Survey of the Second SEEDS Public Workshop

Note: where necessary for readability, grammar and spelling have been fixed. Otherwise, all comments are verbatim.

About the content

1. What was the most useful element of this workshop to you? Why?

   - Reagan Moore's presentation – coordinated overview of work in digital library, grid, and persistent archive communities
   - Structure/format pretty good; provided nice running status of deliberations
   - The talk before the breakouts about breakouts; all parts will be part of the whole and knowing what is up with each is important; Bruce Caron had a great presentation and discussion – would have liked it earlier
   - The homework the SEEDS team was forced to do for this 'deadline'; viewing and commenting on the current state of SEEDS plans
   - Interacting with all participants. It helps vet ideas and establish common ground. Personal relationships also help facilitate future interactions.
   - Actually – the whole thing was a good balance of progress reporting, new info (e.g., the data grid) and breakout discussions
   - Bringing Reagan Moore will help us bridge to NSDL activities
   - The breakout sessions were the most useful element to me. I am a new member to this community of professionals, and the breakouts help me to see the expertise and the issues that exist here as well.
   - Introduction to NASA processes and methods to involve community – get sense of NASA/Goddard science/direction
   - Peer networking. Access to ides and concerns of other organizations and disciplines. Hearing about future plans.
   - Discussions with the different perspectives
   - The life cycle discussion was useful because I had not heard of the plans for transition of data from EOSDIS to NOAA.
   - Learning what SEEDS is about, awareness of issues I had not thought about, and the ability to express my thoughts on issues.
   - The plenary/debrief sessions. These sessions are invaluable for obtaining insight/understanding of issues being raised in fora one individual is unable to participate [in] due to the interest in another session.
   - Since this is the first workshop attended, to me it was very useful in understanding what SEEDS is all about. I will not say that I am going with a clear picture of SEEDS, however this is a start.
   - This was a very balanced workshop with a good mix of presentations and working sessions. Also a good mix of SEEDS and "community" presenters. Better structured and more interesting than first workshop.
   - Capability/technology – this session pointed out where the weak points of the systems are and what improvements are needed.

- Finally got a good understanding of SEEDS and its activities.
- Continuing the dialog on data stewardship, governance and standards. These are critical to the success of NASA data management in the future.
- Community involvement because it needed to be addressed
- Standards – on study team for standards
- Sharing ideas; extracting lists of 'new ideas' for data system consideration; discussion on life cycle spurred ideas on permanent archive; combination of group and breakout sessions
- There were some excellent ideas and discussions from the capability/technology breakout session
- Discussion of software reuse for mission critical systems; useful presentation on Federation activities
- Breakouts – especially standards breakout; opportunities for interactions with colleagues – breaks, lunches together, breakfasts; extensive intros that set the stage for workshops; Townshend and Caron's talks
- Breakouts – enabled discussion and debate; community involvement
- Different points of view represented; feedback on standards study
- Format pretty successful (i.e., plenary/overviews, working session, report out); invited speakers/perspectives added a lot
- Standards – good presentation on lessons learned by other organizations
- Reuse – very close to current work and also has possibilities to save ESE time and costs
- The MODAPS pitch on reuse and the reuse session. It was the most interesting and let me think about future missions.

2. What was the least useful element of this workshop to you? Why?

- Concurrent sessions – overlapping discussions
- None. I found the information all helpful to the "big picture"
- Some of the freewheeling 'discussion'; some of the emphasis on technology solutions
- Content was fine. Logistics: don't like to have to travel on Sunday. I get little enough time with my family.
- Breakouts need to also break out into discussions of 6 – 8 people; the breakouts were too big.
- Focus on data access rather than data use and value; simple provider/user segmentation rather than production chain or loop where each entity is both.
- At this point, I really can't say. I have not had the exposure to the community and its politics to really make a truly educated opinion.
- Community discussion – I think it needed to be said and was quite good – but the ESE has an established community.
- Mulling through old issues that have no ownership or plan of attack. It creates frustration in the group.
- The discussion of NASA internal policies, contracting, etc.

- All were useful, but it wasn't clear that our breakout groups added value to what had been done previously.
- I found the workshop as a whole to be very useful to me.
- Integration: Sometime soon will need to integrate formulation teams. Okay to leave separate, but need to integrate discussions of 2 or more teams that rely on discussion of other teams – melt the bricks. Why? To assure sound foundations are being formed.
- I believe every element is useful. However a brief summary of SEEDS for a 10-15 minute period would have set the tone better for the workshop. I know that people always can read thru web pages, nevertheless this is a suggestion.
- High cost of group dinner discouraged many from attending. This was peripheral to the business of the workshop. But it's a failure as a "community building exercise" if you price most people out.
- Standards. There is a need for standards but I think that SEEDS needs to get more input from users and their problems with obtaining data.
- Some breakout sessions were not well focused – did not deal with concrete things (too abstract and theoretical)
- Community engagement. We need to recognize the fact that the scientific community does not, nor should they, care about the design and operation of the data systems. They care about quality data and services. You cannot engage them in systems discussions or standards issues. The tools and services should mask the standards.
- The lengthy plenary briefings. Perhaps they should be left for the intro of each breakout.
- Community engagement – needed more up front analysis of what is the relevant community for each SEEDS activity (SEEDS reuse community is different from SEEDS user activities, etc.)
- Interaction with the end-user community as it only included John Townshend.
- Sorry, nothing comes to mind
- Data grid technology to manage data
- Not enough depth – need more end user representation
- Anything having to do with 'creating community involvement' – methodologies for getting deep community involvement. Don't know scope of SEEDS, therefore don't know who their customer is. Lots of talk, but just focusing on ESE responses
- Standards, because typically they are set or specified for me

3. What topics do you wish we had covered, and who might talk about them?

- Tested processes – someone from OGC, IETF, etc.
- NSDI/GSDI development – someone from FGCD
- Future remote sensing missions over 5 – 15 years – NASA, NOAA, commercial

- HQ discussion on how the SEEDS process will work. Give examples. What near-term missions will be affected by SEEDS? Explain how NPP will be affected by SEEDS.
- Don't know what I don't know
- An overall view of the data 'policy' at NASA – SEEDS is only trying to facilitate interfaces, standards, etc – since <u>data</u> flows through those interfaces, understanding NASA Y's vision for overall data productions, distribution, etc., is essential.
- I'd like to learn more about HQ view, funding, and outreach.
- Evolving corporate culture, communication, community within NASA ESE…some experts on corporate culture might be found.
- BENEFITS! Can't separate LOS or cost from discussion of value (with metrics that reflect value). Can't do strategy until goals are linked to their expected value to each population.
- I think a topic that was evident from the comments presented over these last few days would be defining scope in relation to SEEDS.
- Current processes in use by successful programs/teams – lessons learned
- Legacy processes worth keeping
- Maybe another success story (like MODIS fire) that has benefited from the SEEDS spirit and intent.
- Lessons learned – issues and resolution tactics of a real world project (EO-1, SRTM on the land science side)
- The nature of future missions and projects is very important, but not known to everyone.
- Unfortunately, I had to leave early and so missed the metrics session, which I would have been very interested in.
- The benefits SEEDS will bring to the Earth Science community.
- Knowledge-based systems – extraction of knowledge from data/info, organization of knowledge for query/results; presentation of knowledge – some <u>corporate</u> expert
- We are developing SEEDS as an interface and data links to be used by the <u>user community</u>. An independent validation and verification as to its performance should be, in my view, a topic to be covered. Maybe we can ask Fairmont/NASA facility to participate.
- How to manage the development of SEEDS in distributed environment.
- More on where the major problems or bottlenecks are in the present system.
- The focus needs to move toward implementation. We cannot plan and formulate forever. How are we going to design the overall infrastructure? How will it be funded in the long term? What will be the governance structures?
- The future interaction between the SEEDS office, ESE, and the implementing programs. People who will be doing the actual work – Steve, Mike Luther or his rep, the NPP or other mission person
- What other space agencies are practicing in all of these topics

- How is SEEDS going to expedite needed studies that impact near-term missions?
- Coverage was adequate
- End-user involvement; applications; COTS providers
- A review of the 1st 5 years of EOSDIS and what was done at that time about many of these same issues. Don't just identify what went wrong, but also what was done well and what went right.
- Perhaps we could get various hardware and software vendors to talk about their plans for the next several years
- Now that governance is addressed…..none

4. What topics do you wish we had left out and why?

- None
- I do not know – it all seems necessary.
- I think all the topics discussed are topics that were needed considering where SEEDS is in its conception/formulation.
- The technology transfer segment wasn't as useful. It wasn't specific. It was good for outreach to the general public. It just didn't offer anything substantial to the group.
- All good.
- None
- No opinion now, but we may want to downplay community involvement in near future.
- None
- Testing. Testing points out the need for strong ICD's, OA's.
- None
- None
- MODAPS reuse – was not presented in a manner that described usefulness to other projects
- Metrics – doesn't add much info at this point
- Community involvement – we need something innovative and different for engaging the science community, but the key SEEDS community is the data provider/data system developer community. They in turn have end users of data as their user community. SEEDS needs outreach to this secondary community.
- Bruce Caron's talk on communities. Need more than just an academic description and to focus on concrete strategies (or lessons learned) to accomplish this vision. John Townshend talk on definitions was trivial and the only useful part was the results of the Google search on the CDC.
- The first implementation talks – they were interesting, but about implementation and not requirements. Update – the Federation talk was less interesting and/or relevant. Now that the Federation has the Foundation, I'm not sure it can be given special treatment by NASA.

5. What topics should we add to the next workshop and who are the experts?

- Security architecture – not sure who
- Direct broadcast – current experience, future trends
- Data center/data management – "life cycle" is not just long-term archiving!
- More on OAIS, benchmark study
- Data integration – cross-disciplinary data use
- Intellectual property issues – Paul Uhler, (NAS)
- See #3
- IBM ESIP marketplace tools – Chung-Sheng Li
- NSDL and links outside NASA – Dave Fulker
- The scope of SEEDS should be addressed, at the very least for those of us who are new to SEEDS, and as a refresher for some fellows here who could use a reminder.
- Knowledge management/sharing – community knowledge system
- Have some users come tell some ESE success stories
- NASA data buy projects (EarthSAT, LDCM); how do these commercial activities fit into SEEDS? Commercial launch segments may take a bigger chunk of new missions.
- It is very important to examine and join forces with the grid activities.
- Issues as seen from the science community as opposed to issues of data management.
- Knowledge-based management – some rep from a corporate organization who is an expert in knowledge-based systems.
- IV&V area people from Fairmont/WVA NASA
- The use of wireless technology and data mining would be useful.
- Would be nice to identify target audience/user of SEEDS data – research scientists (instrument and multidiscipline), modelers, applications, other agencies. Perhaps we could hear from some of them….
- It would seem that SEEDS main function is to tie together the various systems. It would be good to see what can be improved or made better in the interfaces and in the web access area.
- Implementation – implementing agencies, i.e., NASA, NOAA, USGS, others – NSF?
- Interaction between SEEDS office, HQ, and programs; until we know that, we don't know if this is wasted effort or not. I don't know who the experts are.
- Federation and SEEDS: How are they going to get along?
- Ease of use of ES data – insights from knowledge discovery/NSF, DLESE; information grid – Ames; ES modeling framework (climatic focus) – NCAR is lead; intellectual property rights; software release process, tech transfer "hoops" – is there anyone who understands and can make [these last three] interesting?; open source or Linux speaker – describe process

- More industry/[can't read] representation – their perspective on standards/trends
- More on DAAC/SIPS perspective
- Something on grids, esp. data grids. Not just a sales job, but something that gets into the pros and cons. A series of user perspectives, especially from users who have lived through the 1st 15 years. Alternative collection management approaches, other than SRB.
- User presentations from oceanography, geology, hydrology, etc., as to their future needs
- Having a "community engagement" workshop may be useful – identify specific communities broken down by each SEEDS activity; discuss how to get representative participation for each activity by each community

6. Please identify 2 people who are not here this time who should be next time, and how we can reach them?

- More NOAA – LTA involvement
- Flight project/instrument team representatives to gather input and requirements from this side of the user community
- Dr. Mike King/MODIS Scientist GSFC/NASA
- Dr. Helen Wood/NOAA NESDIS
- Dr. Dan Tarpley/NOAA NESDIS ORA
- Rep from NCSA (HDF expert), although here this session, should be invited to speak about the future of HDF (HDF EOS) – long term directions and support for this technology
- NASA program manager
- Representatives from the science community – offer travel money to support them and make them aware of the importance of their attendance and participation in the group
- Ron Birk (or other) to get buy-in – HQ – need to have HQ support in addition to Martha Maiden
- PMs from NASA projects to get buy-in
- Users and new blood from outside the same old bunch that have been together for 10+ years- for ideas
- Users – pay them to come – via PMO
- I am here representing The Numerical Terradynamic Simulation Group, we are part of the science community, I think more members of the Earth Science community should be here.
- Chung Sheng Li at IBM Watson
- Dave Fulker at UCAR
- Dave Fulker – UCAR NSDL PI
- Dave Mogk – DLESE, Montana State University
- Dr. David Glover

- We need more people to come from ASF, one person can't be everywhere. Most everyone else came as a group (or at least a few)
- If users are important, we should contact all our users and invite them, but then there should be plenty of time for their input
- Cynthia Rosenzweig, GISS – on <u>ongoing</u> climate change
- Someone from new Climate Change Office (Mahoney?)
- Ron Birk et al, NASA Applications
- (good question) In general, each DAAC, ESISS, NRC, NSF, ESIPs, SIPS, and other organizations should be asked to send 2 users representatives (e.g., DAAC user working groups, members of ESISS, NRC)
- We should have users of various disciplines speak to their requirements for STDs, reuse, metrics, etc.; science community people such as Nick VanDriel and Charlie Trautwein
- Dr. Chris Justice at UMD
- Charlie Catlett of the "global grid forum'; has ES background but can speak very knowledgeably on grids and also can speak knowledgeably about processes such as IETF and GGF, both of which are models for the SEEDS standards processes. David Stern, formerly of RSI (IDL) to speak on vendor perspectives. Don Sawyer and/or Lou Reich to talk on their archiving work, and also on related topics.
- Instrument team representative – talk about experience, reuse, etc.
- Joy Henegar (GSFC/NPP); Dan DeVito (GSFC/NPP/IPO)
- 2 more science folks – a NASA person like Jim Irons, and some non-NASA scientist

7. Are you willing to actively help the SEEDS Formulation Team with present and future activities? If so, please give us your specialty and your coordinates so we can reach you!

- Bob Chen – user community interactions; standards development; data integration; DAAC issues
- Victor Zlotnicki – ocean science; DAAC scientist
- I'm on a study team…
- Bruce Caron – issues of community development
- Howard Burrows – sure
- Michael Singer – contractor – data system development
- Tim Smith – long term archiving of land science data (NSLRSDA)
- Bob McGrath – I would be interested in the technology infusion study
- Mary Armstrong – I'd be very interested in areas of LOS, metrics, reuse
- Sherry Harrison – I am involved with user services, data distribution and operations at the Global Hydrology Resource Center. I am also currently chair of the DAAC Alliance User Services Working Group (USWG).
- Vince Troisi – maybe; engineering management
- Sarma Modali – development of user community
- James Fischer – IV&V, testing, performance of systems

- I already am
- Yes. My specialty is data models and file formats, esp. HDF, and also how these interact with high performance computing (Mike Folk)
- Mission critical data systems (Ed Masuoka)
- I am currently helping Matt Schwaller with the Life Cycle study
- Life cycle (data stewardship); less active, but participate in LOS and technology (Steve Kempler)
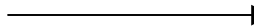- Yes – Yonsook Enloe

8. Other comments?

- I hate to suggest more/longer meetings, but study teams need to interact more. Lots of convergence – e.g., Ed's talk on reuse was full of examples of benefits of standards (data model, HDF EOS) and interfaces (modular, plug-n-play design)
- Standards and lifecycle groups have identified various resource requirements that should feed into cost model.
- Need more inputs from the user community
- If the SEEDS workshop holds in the same form as before, NASA should advertise the workshop in the major professional newsletter (e.g. EOS) to invite the participants
- SEEDS formulation team should hold workshop in conjunction with major professional meetings to solicit inputs from user community
- Wish we had approached other agencies. USDA (has a remote sensing group that uses satellite data), DOE (has at least one major program ARP), EUMETSAT and ESA, ISRO/INSAT group/NRSA group – these people learned to use data from US satellites
- It would be very helpful to have the status reports from the working groups before the workshop, rather than relying on a 30-minute summary.
- Great venue (San Diego)
- Good organization and coordination (good job Kortney)
- Provide presentations on web site
- Great accommodations/great food
- Pick a hotel with broadband access please!!
- I feel this was a better workshop than last time. I could more easily participate and I learned far more. I would like to have had the question sheets before I came so I could talk with those who didn't come but would have valuable input.
- When people ask what SEEDS is, it would be nice if there were a standard answer. Everyone seems to have a different twist. [can't decipher remainder of the comment]
- I think the lack of FT study team cross-information flow is hurting the FT effort.

- Group dinner too expensive; Starting Monday morning after Father's Day was very family unfriendly; Good weather! Needed more beach time…
- Very good workshop, especially in extracting new ideas. Has there been consideration for developing a distributed data system model that includes the ideas of various breakout sessions? This would validate topic ideas and help develop processes. This would also provide a higher look at all the areas of study working together.
- Perhaps we could do marketing to gain user support. Advertise in PS&RS and other periodicals and attend user conferences.
- Be sure and set up a web-based or e-mail based method of marking-up existing SEEDS documents and adding feedback from end-users.
- Presentations frequently were really hard to read because the font was small and/or there was way too much text, and inevitably the speaker didn't provide sufficient time to read the text. How about advising speakers on how to do good presentations. There are some good sources for this. Provide an e-mail room. Wireless access would be nice but isn't critical. Kortney Stevens is a winner!
- Focus on integration of study team findings in next SEEDS workshop.
- Should have spent a lot more time on orientation and background of SEEDS, and especially provide this info on handouts for future reference (at the breakout session).
- Interesting and necessary groundwork.

About the breakout sessions

Please rate the breakout sessions. In terms of your work and your interests, how useful were the breakout sessions to you.

	Least useful ————————▶ Most useful				
Community Involvement	1	2	3	4	5
18 responses			3.61		
Capability/Technology	1	2	3	4	5
25 responses			3.68		
Standards	1	2	3	4	5
22 responses				4.30	
Life Cycle	1	2	3	4	5
17 responses				4.08	
Reuse	1	2	3	4	5
17 responses			3.59		
Metrics	1	2	3	4	5
14 responses			3.36		
Level of Service/Cost Model	1	2	3	4	5
12 responses			3.50		

Comments on the breakout sessions?

- Substantial reliance on surveys that were not conducted with basic controls on bias, sampling, reliability, human subjects protections. Qualitative vs. quantitative use; former may be justified, but latter is not. Difference between surveying an entire population and a sample of a population.
- Too many were unfocused – lots of unfocused discussion; material covered in the past; "policy"/guidelines had to be rehashed. I feel that few visited and read the material from the SEEDS #2 website. Maybe people should be pre-assigned (volunteer) to a particular breakout session and let them prepare for it using web material and other materials provided by the chair. Folks not pre-assigned are free to attend any session.
- Life cycle presentations were great starts; wish we had more time to follow thru.; cap/tech – could have split earlier; Bruce's talk – outstanding; the whole meeting much better/ in group there was the hostility and hard headedness. People [rallied?] and those that weren't in the core group could participate.
- Sessions were good. Summaries were also good.
- I would ask how useful were the breakout sessions for SEEDS?
- Use [can't read word] to capture and expand concepts.
- There seems to be some angry undercurrents here. I feel that it has actually undermined, to some extent, the insight and creativity of a couple of the breakouts that I attended. The comments and bickering took up valuable time. These issues should be handled in another forum. This would free us up to solve the issues at hand and save the angst fro a more appropriate forum.
- Good forum – the approach used – i.e., give introduction and continue in breakout sessions is very good and worked well. Before I came here, I did not understand the agenda – maybe you could explain it for SEEDS #3.
- They were very useful. Breakout sessions are a great way to share ideas and learn about the needs of others involved with Earth science data at various levels.
- Overall very educational and interactive! Eventually we'll need an "uber" breakout session that may last a bit longer but does not conflict with other sessions. The plenary provides some opportunity for joint interaction, but I believe a session joining say standards, reuse, and capability/technology will be needed soon.
- They were useful as you hear from various individuals their opinions.
- Breakout sessions are most useful for the formulation team to get inputs. In the future, workshop should reduce the number of plenary sessions and increase breakout sessions. Move some of plenary talks to breakout sessions.
- I'm generally interested in more breakouts than I can attend. Then again, 3 days is about the maximum amount of time you can sustain interest in a workshop – maybe that's why I was interested in all 4 Mon/Tues breakouts, less interested in Weds breakout. A more useful comment: breakout timeframes were about right.
- Most of the sessions were useful. Somehow a need to get more users of the data involved in SEEDS is needed. A survey using the database of users from EDG may be a way to inform users of SEEDS and also to get input on what they would find helpful in tools or we access.

- External speakers generally useful; general format of workshop worked well, but some breakouts needed more time; near term stds and long term stds process should have been separated into own breakouts – ran out of discussion time; at some point, having a 2 – 3 day meeting on only one study topic, with invited targeted reps, representing different communities to work out topic or tire kick process, etc., may make sense and be useful
- It's good to have presentation in breakouts, but important to make these brief enough to leave plenty of time for discussion; give handouts of key elements of presentations, especially flow diagrams that are hard to remember; moderators did an excellent job of keeping focused and were really well prepared – Great!; summaries also well done, but make sure these are on the web.
- Good forums for brainstorming for the SEEDS team.
- Technology – not enough free form; too question focused; leading to known answers
- Important to provide rapid feedback on at least preliminary results via SEEDS web site by 6/24

Your interest in SEEDS

- Data stewardship
- Need to know what is coming in the future
- Community development; interaction with science user
- Earth science data users access to data and customer service
- Understanding and participation in development/evolution of next generation systems
- Lifecycle/LTA
- Data systems
- The standards and technology understanding and expertise could really help us accomplish our goals.
- NASA name and EOSDIS is inspiring and exciting to the general public. Can use this to improve science practices generally.
- Helping promote democratic ownership within the NASA/user community partnership.
- General as ASF will fit in and how ASF will be able to participate
- DAAC Alliance vice chair; Federation ESIP-1 rep; involved in distributed web mapping, standards development
- NPP SDS
- I expect to work on NASA missions in the SEEDS era
- Its concern with the balance between community involvement and mission requirements; standards evolution; archiving
- Near-term standards
- Reuse
- On long-term standards process study team
- Facilitate the evolution of data systems to support Earth science, specifically, preservation of data and related levels of service; technologies
- Member of the life cycle study team
- Developing low-cost data systems; new services provided

Where should we hold our next workshop?

- East (northeast?)
- Maryland
- No preference
- Not College Park
- Somewhere the weather is warm and the fishing is good! ☺ Seriously, though, possibly Boulder (start hosting meetings where we can directly involve community members).
- Away from the office was good – got good NASA attendance
- A hub city – like St. Louis or Minneapolis; mid US; Denver/Boulder – NOAA hosted?
- Middle of US

- Location is not really important
- NCSA – Champaign, Ill.
- Maryland
- San Francisco in conjunction with AGU fall meeting
- East Coast, not Greenbelt/College Park. Prefer city (like Bethesda) with many things in walking distance, or resort area (ocean or mountains) to generic hotel. But network availability important.
- Ashville, maybe tour the NOAA archive
- Hawaii; JPL/GSFC because of the instrument teams and flight projects
- Maryland/Washington/Virginia (maybe Eastern Shore)
- Colorado
- Cancun, Mexico
- Chicago, Denver, St. Louis, or Mpls/St. Paul, centrally located hub cities – or Champaign, IL for NCSA
- Middle of the continental United States
- Boulder/NCAR; Nashville (?)/NCDC
- In Maryland or Virginia, somewhere not too far from Goddard Space Flight Center